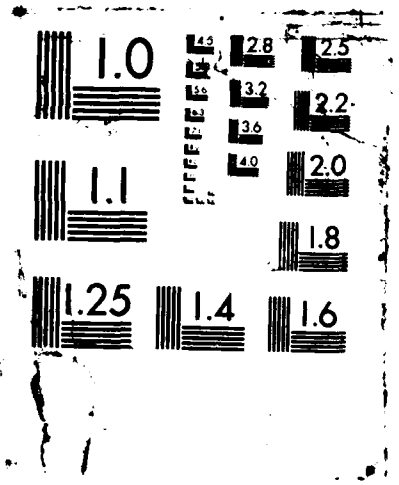
AD-A191 442 COMPUTER-AIDED DESIGN OF THERMOSTABLE PROTEINS(U) JOHNS 1/1
HOPKINS UNIV BALTIMORE MD DEPT OF MOLECULAR BIOLOGY AND
GENETICS C O PABO 31 MAR 88 5L406 N00014-86-K-0196

UNCLASSIFIED F/G 6/1 NL


1.0
2.8
2.5
3.2
2.2
3.6
4.0
2.0
1.1
1.8
1.25
1.4
1.6

DTIC FILE COPY

(2)

SECURITY CLASSIFICATION OF THIS PAGE

AD-A191 442

REPORT DOCUMENTATION PAGE

	1b RESTRICTIVE MARKINGS NA
2a. SECURITY CLASSIFICATION AUTHORITY NA	3. DISTRIBUTION / AVAILABILITY OF REPORT Distribution unlimited
2b. DECLASSIFICATION / DOWNGRADING SCHEDULE NA	
4. PERFORMING ORGANIZATION REPORT NUMBER(S) 5L406	5. MONITORING ORGANIZATION REPORT NUMBER(S) NA
6a. NAME OF PERFORMING ORGANIZATION Johns Hopkins University 6b OFFICE SYMBOL (If applicable) NA	7a NAME OF MONITORING ORGANIZATION Office of Naval Research
6c. ADDRESS (City, State, and ZIP Code) Department of Molecular Biology & Genetics 725 North Wolfe Street Baltimore, MD 21205	7b ADDRESS (City, State, and ZIP Code) 800 N. Quincy Street Arlington, VA 2217-5000
8a. NAME OF FUNDING / SPONSORING ORGANIZATION Office of Naval Research 8b. OFFICE SYMBOL (If applicable) ONR	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER N00014-86-k-0196
8c. ADDRESS (City, State, and ZIP Code) 800 N.Quincy Street Arlington, VA 22217-5000	10. SOURCE OF FUNDING NUMBERS PROGRAM ELEMENT NO. 61153N; PROJECT NO. RR04106; TASK NO. NR441g003; WORK UNIT ACCESSION NO

11. TITLE (Include Security Classification)

Computer-Aided Design of Thermostable Proteins (U)

12. PERSONAL AUTHOR(S)
Carl O. Pabo

13a. TYPE OF REPORT	13b. TIME COVERED	14. DATE OF REPORT (Year, Month, Day)	15. PAGE COUNT
Annual	FROM 03/15/87 TO 03/14/88	March 31, 1988	5

16. SUPPLEMENTARY NOTATION

17. COSATI CODES			18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)
FIELD	GROUP	SUB-GROUP	protein design/thermostability directed mutagenesis/protein engineering

19. ABSTRACT (Continue on reverse if necessary and identify by block number)

See reverse side.

DTIC ELECTE S APR 06 1988 D H

DISTRIBUTION STATEMENT A
Approved for public release;
Distribution Unlimited

20. DISTRIBUTION / AVAILABILITY OF ABSTRACT	21 ABSTRACT SECURITY CLASSIFICATION
☒ UNCLASSIFIED/UNLIMITED ☐ SAME AS RPT ☐ DTIC USERS	U

22a NAME OF RESPONSIBLE INDIVIDUAL	22b TELEPHONE (Include Area Code)	22c OFFICE SYMBOL
Dr. Michael T. Marron	(202)696-4038	ONR

DD FORM 1473, 84 MAR 83 APR edition may be used until exhausted
All other editions are obsolete

SECURITY CLASSIFICATION OF THIS PAGE

We are developing strategies for computer-aided protein design. Our strategies emphasize simple geometric aspects of protein structure, and our computer program (PDB_PROTEUS) allows us to systematically test a very large number of alternative sequences and conformations.

We have written programs to systemtically search a protein structure for places where we can add disulfide bonds, new salt bridges, or favorable aromatic interactions. We also search for places where the backbone can accommodate glycine ->alanine mutations and for places where prolines can be accommodated. (These changes -like the disulfide bond- should stabilize the folded structure by reducing the entropy of the unfolded form.)

We have tested each of these approaches by trying to design thermostable variants of the lambda repressor. The most stable variant - ala46ala4cys88lys93 - contains 4 mutations and is 17 degrees more stable than the wild-type protein. Since mutations that reduce the entropy of the unfolded form seem to have the largest stabilizing effect on repressor, we are searching for ways to introduce new metal-binding sites and chemical crosslinks between side chains.

Computer-Aided Design of Thermostable Proteins

We are developing methods for computer-aided protein design and are testing these strategies by constructing thermostable variants of the lambda repressor. Repressor's DNA-binding domain normally denatures at 54° C; out goal is to construct a variant that will bind normally to DNA and yet be stable to 100° C. To date, we have constructed a quadruple mutant that is stable to 71° C and binds DNA as well as the wild type protein.

Our fundamental goal is to develop methods for de novo protein design, and we are proceeding by treating the problem of protein design as an "inverted" version of the protein folding problem (Pabo, 1983). In the protein folding problem, one is given an amino acid sequence and must predict how this folds in three dimensions. Protein design can be approached in quite a different way - one can begin by chosing a folded arrangement of the polypeptide backbone and then try to pick an amino acid sequence that will stabilize this structure. "Inversion" eliminates the problem of predicting long-range interactions, since residues which will interact in the final tertiary or quaternary structure already are close in space when they are added to the prefolded backbone. One should be able to pick residues which will have favorable interactions with their neighbors.

We are developing a program, called PDB_PROTEUS, for computer-aided protein design. Our program uses simple geometric aspects of protein structure and frequently uses local coordinate systems so that the geometric relationships are easier to visualize (Pabo and Suchanek, 1986). There are many advantages to using a computer program: A program can easily check millions of possible sequences and conformations. Using a program also makes it easy to try several variations of a particular search strategy or to apply the same strategy to many differrent proteins. Keywords: Computer-aided design; molecular biology; computer programs; synthesis chemistry; computer simulation

Much of our work during the past year has focussed on developing and refining the PDB_PROTEUS system. Although the program is written in FORTRAN, we have tried to develop a programming strategy that will be very flexible. The core of the system is a library of subroutines. Each performs a discrete

operation - like adding a residue or changing the coordinate system - and our library now contains about 120 subroutines. The main programs use these subroutines (almost like a higher level programming language). Since we believe these subroutines and programs could be useful for many other projects, we have paid careful attention to our programming style and are making this system available to other laboratories.

In our attempts to stabilize the repressor, we have written programs to seach the repressor structure for the best residues to change and then have experimentally tested each of these predictions. We have tried five different types of changes, and our results are summarized below:

1) Disulfide Bonds: The search program used all the disulfide bond conformations found in the Protein Data Bank and also used a library of conformations that were closely related to the left handed spiral configuration (Richardson, 1981). Modelling suggested that an intermolecular disulfide bond could be introduced by changing Tyr 88 to Cys (Pabo and Suchanek, 1986). Experimental studies show that this disulfide bond forms spontaneously, stabilizes the repressor dimer against thermal denaturation, and increases the affinity for DNA (Sauer *et. al.*, 1986).

2) Salt Bridges: The program searches for any position where a new salt bridge could be introduced by changing a single residue. The best position appeared to be at the C-terminal end of helix 5, where changing Ser 92 to Lys should allow a salt bridge with Glu 89. This has no effect on the thermal stability of repressor, but introducing Lys 93 (effectively adding a residue to the C-terminal end of the helix!) does stabilize the protein by about 0.5^{o} C. Although this approach does not seem to be very promising, we need to test additional positions.

3) Aromatic Interactions: Studies of aromatic-aromatic interactions in proteins suggest that these can stabilize a protein if the aromatic rings are about 5.5 A apart and are approximately perpendicular to each other (Burley and Petsko, 1985). We have searched for places were aromatic residues could be added to make favorable contacts with an existing aromatic residue.

Unfortunately, the only position that appears plausible (residue 33) changes a key residue involved in nonspecific contacts with the DNA. Studies in Robert Sauer's laboratory at M.I.T. have shown that this mutant is stable (Hecht *et. al.*, 1984) but it is not useful to us because it disrupts DNA binding.

4) Glycine to Alanine Changes: Hecht and Sauer (1986) have shown that repressor can be stabilized by changing two glycines to alanines. We have set up a program that automatically searches for places that Gly to Ala changes might be made. This does not find any other plausible postions in repressor, but the program should be useful with other proteins.

5) Prolines: Proline residues may stabilize proteins by reducing the conformational entropy of the unfolded protein. Obviously, they can introduce unfavorable strain if they are put at the wrong postions, but we have search repressor for positions where the backbone conformation should allow a proline residue to be introduced. Two positions appeared plausible and have been tested. We found that changing Tyr 60 to Pro has a mild destabilizing effect, but changing Gln 9 to Pro stabilizes the protein by 0.6^{o} C.

Since we expect tnat a set of changes will be necessary to stabilize repressor to 100^{o} C, it was important to determine whether the effects of multiple mutations are additive. Our initial results are quite encouraging. To test the effects of multiple mutations, we combined our disulfide mutant with the two glycine to alanine changes in helix 3. We found that the wild type protein denatured at 54^{o}, the Cys 88 mutant denatured at 62^{o}, the Ala46Ala48 double mutant denatured at 62^{o}, and the Ala46Ala48Cys88 mutant was stable to 70^{o} (Stearman *et. al.*, 1988). More recently, we have shown that that the Ala46Ala48Cys88Lys93 quadrupule mutant is stable to 71^{o}. We are now introducing the Pro 9 mutation to see whether this gives a further incremental increase in the stability.

During the last year we also have continued with the crystallographic refinement of repressor, and it is clear that a highly refined structure is very important for modelling and design. Our initial predictions had used the repressor structure obtained by fitting to an isomorphous electron density map at 3.2 A

resolution (Pabo and Lewis, 1982). We have much better data from our repressor-operator cocrystals (Jordan *et. al.*, 1985) and this structure has been refined to an R factor of 24.5% using data from 8.0 to 2.5 A resolution. Comparisions have shown that our model-building predictions are very sensitive to differences between these coordinate sets. The intial, less accurate, coordinates gave several predictions (not discussed above because they were not obtained with the better coordinates) that were thermally unstable.

Although it is too early for a firm conclusion, our data suggest that it may be easiest to stabilize a protein by introducing mutations that reduce the entropy of the unfolded form, and we will focus on this strategy during the final year of our contract. Our disulfide search program can readily be rewritten to search for any types of crosslinks between side chains. (The program only requires that we can predict plausible conformations for the crosslinks, because we do need to know the relative spatial arrangement of the proximal and distal residues that are connected by the crosslink.) Specifically, we plan to introduce cysteines that can be connected by reagents like dimercaptoethanol, dithiothreitol, or can be bridged by a mercury ion. (Many other crosslinking reagents can be used, but we think it will be advantageous to begin with ones that react reversibly, since it should be easier to analyze the products.) These approaches may allow us to reach our goal of designing a repressor variant that is stable to 100^{o} C.

References

Burley, S.K. and Petsko, G.A. (1985) Aromatic-Aromatic Interaction: A Mechanism of Protein Structure Stabilization **Science 229**, 23-28.

Hecht, M.H., Sturtevant, J.M., and Sauer, R.T. (1984) Effects of Single Amino Acid Replacements on the Thermal Stability of the NH2-terminal Domain of Phage Lambda Repressor **Proc. Natl. Acad. Sci. USA 81**, 5685-5689.

Hecht, M.H., Sturtevant, J.M., and Sauer, R.T. (1986) Stabilization of Lambda Repressor Against Thermal Denaturation by Site-Directed Gly to Ala Changes in Helix 3 **Proteins 1**, 43-46.

Jordan, S.R., Whitcombe, T.V., Berg, J.M., and Pabo, C.O. (1985) Systematic Variation in DNA Length Yields Highly Ordered Repressor-Operator Cocrystals **Science 230**, 1383-1385.

Pabo, C.O. (1983) Molecular Technology: Designing Peptides and Proteins **Nature 301**, 200.

Pabo, C.O. and Lewis, M. (1982) The Operator-binding Domain of Lambda Repressor: Structure and DNA Recognition **Nature 298**, 443-447.

Pabo, C.O. and Suchanek, E.G. (1986) Computer-Aided Model-Building Strategies for Protein Design **Biochemistry 25**, 5987-5991.

Richardson, J. (1981) Protein Anatomy and Taxonomy **Advances Protein Chemistry 34**, 167-339.

Sauer, R.T., Hehir, K., Stearman, R.S., Weiss, M.A., Jeitler-Nilsson, A., Suchanek, E.G., and Pabo, C.O. (1986) An Engineered Intersubunit Disulfide Enhances the Stability and DNA-Binding of the N-terminal Domain of Lambda Repressor **Biochemistry 25**, 5992-5998.

Stearman, R.S., Frankel, A.D., Freire, E., Liu, B., and Pabo, C.O. (1988) Combining Thermostable Mutations Increases the Stability of Lambda Repressor, **Biochemistry**, submitted

AMZEL, L. Mario
Department of Biophysics
Johns Hopkins School of Medicine
725 North Wolfe Street
Baltimore, MD 21205

BEVERIDGE, David L
Department of Chemistry
Wesleyan University
Hall-Altwater Laboratories
Middletown, CT 06457

FICHT, Allison R.
Department of Medical
Biochemistry & Genetics
Texas A&M University
College Station, TX 77843

FRAUENFELDER, Hans
Department of Physics
University of Illinois
Urbana, IL 61801

GABER, Bruce P.
Code 6190
Bio/Molecular Engineering Branch
Naval Research Laboratory
Washington, DC 20375

HONIG, Barry
Columbia University
Dept of Biochemistry and
Molecular Biophysics
630 West 168th St.
New York, NY 10032

LAURSEN, Richard
Chemistry Department
Boston University
590 Commonwealth Avenue
Boston, MA 02215

LENZ, Robert W.
Chemical Engineering Department
University of Massachusetts
Amherst, MA 01003

LEWIS, Randolf V.
Biochemistry Department
University of Wyoming
University Station Box 3944
Laramie, WY 82071

LINDSAY, Stuart M.
Department of Physics
Arizona State University
Temp, AZ 85278

McCONNELL, Harden M.
Stanford University
Department of Chemistry
Stanford, CA 94305

MEYER, Edgard F.
Texas A&M University
Dept of Biochemistry and
Biophysics, Box 3578
College Station, TX 77843

NAGUMO, Mark
Bio/Molecular Engineering Branch
Code 6190
Naval Research Laboratory
Washington, DC 20375-5000

NOVOTNY, Jiri
Laboratory of Cellular and
Molecular Research
Massachusetts General Hospital
Boston, MA 02114

PABO, Carl O.
Department of Biophysics
Johns Hopkins University
School of Medicine
Baltimore, MD 21205

PRENDERGAST, Franklyn
Mayo foundation
200 first st. SW
Rochester, MN 55905

RICH, Alexander
Department of Biology
Massachusetts Institute of
Technology
Cambridge, MA 02139

RICHARDS, J. H.
California Institute of
Technology
Division of Chemistry and
Chemical Engineering
Pasadena, CA 91125

SHULER, Michael
School of Chemical Engineering
Cornell University
Ithaca, New York 14853

SINSKEY, Anthony J.
Laboratory of Appl Microbiology
Dept of Appl Biological Sciences
MIT
Cambridge, MA 02139

URRY, Dan W.
Laboratory of Molec Biophysics
University of Alabama
P. O. Box 311
Birmingham, AL 35294

WAITE, J. Herbert
College of Marine Studies
University of Delaware
Lewes, DE 19958

WARSHEL, Arieh
Department of Chemistry
University of Southern California
University Park
Los Angeles, CA 90089-0482

WILLIAMS, Jon I.
Allied Corporation
Columbia Rd and Park Ave.
Morristown, NJ 07960

CHANG, Eddie L.
Bio/Molecular Engineering
Code 6190
Naval Research Laboratory
Washington, D.C. 20375-5000

CORDINGLEY, John S.
Department of Molecular Biology
University of Wyoming
Box 3944 University Station
Laramie, WY 82071

CASE, Steven T.
Department of Biochemistry
University of Mississippi Medical
Center
2500 North State Street
Jackson, MS 39216-4505

Dr. Margo G. Haygood
Office of Naval Research
800 North Quincy Street
Arlington, VA 22217-5000

Administrator (2 copies, Enclose DTIC Form 50)
Defense Technical Information Center
Building 5, Cameron Station
Alexandria, VA 22314

Dr. M. Marron
Code 1141MB
Office of Naval Research
800 North Quincy Street
Arlington VA 22217-5000

ANNUAL AND FINAL REPORTS ONLY (One copy each)
Commander
Chemical and Biological Sciences Division
Army Research Office
P. O. Box 12211
Research Triangle Park, NC 27709

Directorate of Life Sciences
Air Force Office of Scientific Research
Bolling Air Force Base
Washington, DC 20332

Chemistry and Atmospheric Sciences Directorate
Air Force Office of Scientific Research
Bolling Air Force Base
Washington, DC 20332

Director
Biotechnology Division
CRDEC
Aberdeen Proving Grounds, MD 21010-5423

Administrative Contracting Officer
ONR Resident Representative
(Address varies - obtain from your business office)

Director, Code 12
Applied Research and Technology Directorate
Office of Naval Research
800 North Quincy Street
Arlington, VA 22217-5000

Director, Code 22
Support Technology Directorate
Office of Naval Technology
800 North Quincy Street
Arlington, VA 22217-5000

Director, Code 112
Environmental Sciences Directorate
Office of Naval Research
800 North Quincy Street
Arlington, VA 22217-5000

Director, Code 113
Chemistry Division
Office of Naval Research
800 North Quincy Street
Arlington, VA 22217-5000

FINAL AND TECHNICAL REPORTS ONLY
Director (6 copies)
Naval Research Laboratory
Attn: Technical Information Division, Code 2627
Washington, DC 20375

END

DATE

FILMED

5-88

DTIC